THE ACCURACY-THROUGH-ORDER AND THE EQUIVALENCE PROPERTIES IN THE ALGEBRAIC APPROXIMANT *

GEORGE A. BAKER, JR.
Theoretical Division, Los Alamos National Laboratory
University of California, Los Alamos, N. M. 87545 USA

Abstract. In addition to the accuracy-through-order requirement that the defining polynomials not all be divisible by z, as required for Padé and integral approximants, there is the further problem of deficiency as pointed out by McInnes. I prove a finite bound on the deficiency and also prove the accuracy-through-order property for algebraic approximants. In addition I prove the equivalence property for algebraic approximants.

Key words: Padé approximant, Algebraic approximant, Accuracy-through-order property, Equivalence property

In the study of Padé approximants, there are a number of basic properties, which are fundamental, and are so well known that they are often taken for granted. The theory of algebraic approximants is not so well developed and some of these properties are not yet established. The theory of both types of approximants is a special case of the theory of Hermite-Padé approximants and starts with a linear, polynomial-defining equation,

$$\sum_{j=0}^{k} P_j(z) f_j(z) = O(z^{s+1}),$$

for the $P_j(z)$'s plus some initial conditions. The approximant involves the solution of

$$\sum_{j=0}^{k} P_j(z) y_j(z) = 0,$$

where the $y_j(z)$'s are related to each other in a manner akin the the way the $f_j(z)$'s are related to each other. The first basic property is to show that as the degree of contact at the origin in the polynomial defining equation increase indefinitely, then so too does the degree of contact between the $f_j(z)$'s and the $y_j(z)$'s. This property is called the accuracy-through-order property. The next basic property is uniqueness. It consists of two parts. First the uniqueness of the approximant polynomials and second the uniqueness of the solution, given the polynomials, for the approximant. Finally there is the problem of equivalence. It is, in the case of

* Work supported in part by the US Department of Energy (contract W-7405-ENG-36) and in part by the NSF (grant INT-90 13977)

the Padé approximant stated as, for a function $g(z)$ which is analytic at the origin,

$$g(z) = \frac{\sum_{j=0}^{l} c_j z^j}{1 + \sum_{j=1}^{m} e_j z^j}$$

if and only if,

$$[L/M]_g = [l/m]_g \quad \forall\, L \geq l,\ M \geq m,$$

where $[\lambda/\mu]_g$ is the Padé approximant to $g(z)$ with the degree of the numerator polynomial equal to at most λ and the degree of the denominator polynomial equal to at most μ. The Padé polynomials satisfy a defining equation with $s = \lambda + \mu$. The problem of uniqueness of the polynomials has been addressed by Baker and Graves-Morris (1990) and of the solution by McInnes (1992). The Padé approximants also have very useful invariance properties with respect to linear fractional transformations of both their arguments and their values. This property is shared by the algebraic approximants (Baker, 1984). So far however, the accuracy-through-order property and the equivalence property for algebraic approximants have not been adequately treated. We will give results on these properties.

The algebraic approximants were introduce by Padé (1892) and an investigation of the quadratic case was made by Shafer (1974).

Definition 1. *Let a function $f(z)$ be given in terms of its Maclaurin expansion. We define the algebraic polynomials Q_{j,m_j} by the accuracy-through-order principle by means of*

$$\sum_{j=-1}^{k} Q_{j,m_j}(z)[f(z)]^{j+1} = O(z^{s+1}), \tag{1}$$

where the $Q_{j,m_j}(z)$ are polynomials of degree at most m_j and s is given by

$$M = \sum_{j=0}^{k}(m_j + 1) - 1, \quad s = M + m_{-1}. \tag{2}$$

We will use the convention that if $m_j = -1$, then $Q_{j,-1}(z) \equiv 0$. The algebraic approximant is denoted by $\langle m_{-1}/m_0, \ldots, m_k\rangle$ and is defined by the solution $y(z)$ of

$$\sum_{j=-1}^{k} Q_{j,m_j}(z)[y(z)]^{j+1} = 0, \quad \text{where} \quad y(0) = f(0). \tag{3}$$

We will always assume that $f(0) \neq 0$. Of course, we can impose additional boundary conditions at $z = 0$, if necessary to break the degeneracy of the possible solutions. See McInnes (1992) for an additional discussion of this point.

In order to complete the definition of these approximants, attention must be paid to the possibility that (1) does not uniquely define the polynomials. The points where $Q_{k,m_k}(z)$ vanishes correspond to singularities. Consequently, in the case where the polynomials are not uniquely determined it would be desirable, in my view, to eliminate any arbitrary or spurious singularities that are introduced by this lack of uniqueness. We complete these definitions by using the minimal polynomials of Baker and Graves-Morris (1990)

DEFINITION 2. *A solution for the algebraic polynomials of type* $(m_{-1},\ldots,m_k)$ *to* $f(z)$ *is called minimal if it is of the lowest degree in the following sense. First there exists no other solution of type* $(m_{-1},\ldots,m_k)$ *for which the actual degree of* Q_{k,m_k} *is smaller. If there exist solutions of type* $(m_{-1},\ldots,m_k)$ *for which* $Q_{k,m_k}\equiv 0$, *then we minimize the degree of* $Q_{k-1,m_{k-1}}, Q_{k-2,m_{k-2}}$, *etc. to find the minimal solution.*

Uniqueness is insured by their arguments. In addition, we will impose the condition that $\vec{Q}_{\vec{m}}(0)\neq\vec{0}$. Baker and Graves-Morris (1994) have proven that we may do so and, although some types $\vec{m}$ may then fail to exist, the table of approximants is as complete, in so far as the derived approximants are concerned, as was the table of approximants derived from the minimal polynomials. Some of the repeats in the table may however have been dropped by this restriction.

A very important property of Padé approximants is that of accuracy-through order. In that case, the requirement that the denominator polynomial not vanish at the origin was sufficient to show that the Padé approximant agreed with the defining series to the same order in z, as the accuracy of the polynomial defining equations. For integral approximants the same results obtain, but for algebraic approximants, as we shall see, the results are roughly true as well, but the situation is a bit more complex. This question has been addressed by Baker and Graves-Morris (1990) for the case of integral approximants, and by McInnes (1992) in the algebraic case.

In order to analyze the problem of accuracy-through-order for the algebraic approximants, we need, following McInnes (1992), to introduce the concept of *deficiency.* While our minimal definition for the algebraic polynomials and the results of Baker and Graves-Morris (1994) insure that we need only look at cases for which $\vec{Q}_{\vec{m}}(0)\neq\vec{0}$, there is another quantity which is important to consider in the algebraic case, *i.e.*, the coefficient of dy/dz in (5) below.

DEFINITION 3. *The deficiency* d *of the algebraic polynomials to* $f(z)$ *is given by*

$$\sum_{j=0}^{k}(j+1)Q_{j,m_j}(z)[f(z)]^j \underset{z\to 0}{\propto} z^d. \tag{4}$$

THEOREM 4. *If the error in the algebraic polynomial defining equation is* $O(z^{\mathcal{M}})$ *and the deficiency of the algebraic polynomials is* d, *then the Maclaurin series for the algebraic approximant differs from the defining series with an error at worst* $O(z^{\mathcal{M}-d})$ *when we specify that* $y^{(j)}(0)=f^{(j)}(0)$ *for* $j=0,\ldots,d$.

PROOF: First, if we differentiate the approximant defining equation (with error displayed for convenience), we get,

$$\left(\sum_{j=0}^{k}(j+1)Q_{j,m_j}(z)[y(z)]^j\right)\frac{dy(z)}{dz}=-\sum_{j=-1}^{k}Q'_{j,m_j}(z)[y(z)]^{j+1}+O(z^{s+t}), \tag{5}$$

where s is defined by (2) and t is the degree of oversatisfaction defined by,

$$\sum_{j=-1}^{k}Q_{j,m_j}(z)[f(z)]^{j+1}=O(z^{s+t+1}), \tag{6}$$

and where the right-hand side is to be of true stated order. Suppose that the deficiency is $d = 0$, then by letting $z \to 0$ in (5) we can compute y_1 from $y_0 = f_0$, where $y(z) = \sum_{j=0}^{\infty} y_j z^j$. If we now compute the coefficient of z^j in (5) we find that it involves y_{j+1} linearly with the coefficient proportional to the limit in (4) which is non-zero in this case. Thus we can solve in a unique recursive manner for all the coefficients y_j. By the defining equation, the f_j are also a solution within the error, so we get $y_j = f_j$ for $j = 0, \ldots s + t$ which establishes the accuracy-through-order results for this case. If the deficiency $d > 0$, then if we examine (5), since it is also satisfied by $f(z)$ as well as $y(z)$, and $f'(0)$ is finite, the left-hand side vanishes like z^d and so too must the right-hand side. Here we have used the initial conditions $y^{(j)}(0) = f^{(j)}(0)$ for $j = 0, \ldots, d$. We next observe that in the defining equation (3) the coefficients of z^j to z^{2j-1} are linear in y_j, since y_j^2 can not appear at lower order that z^{2j}, as every y_j appears only in the combination $y_j z^j$ so that y_j^2 carries at least a factor of z^{2j}, *etc.* Consider the case of deficiency d. The coefficient of z^{2d+1} must be linear in $y_{d+1}, \ldots, y_{2d+1}$. The coefficient in (3) of $z^{2d+1} y_{d+n}$ for $n = 1, \ldots, d+1$ can be seen to be the coefficient of z^{d+1-n} in the coefficient of dy/dz on the left-hand side of (5). This result follows by noting that this coefficient is the partial derivative of (3) with respect to $y(z)$ and so will give the coefficient of the linear terms. Thus because the deficiency is d, only y_{d+1} has a non-zero coefficient. This coefficient necessarily involves only $y_0, \ldots, y_d$, which we have set equal to the corresponding f_j's by the initial conditions. As we examine the higher orders we find that the coefficient of z^{d+n} is linear in y_n with a non-zero coefficient for $n \geq d+1$, and that, as in the case $n = d+1$ just discussed, the coefficients in this term of y_{n+l} all vanish for any $l > 0$. Thus we may, in a unique and recursive manner, compute all the y_j for $j = d+1, d+2, \ldots, s+t-d$. Since as we have remarked, the f_j's satisfy the equation, we have proven the theorem, unless $d > s+t$. But it can not be that $d > s$, since we have a minimal solution. For if it were so, then the quantity in (4) multiplied by $f(z)$, would be a different solution set (of the same order and the same polynomial degrees) to the polynomial defining equations. A linear combination of these two polynomial sets could then be formed which would give a new solution set of order $k-1$ instead of k, which is a contradiction to our selection of a minimal set of algebraic polynomials. ∎

In order for algebraic approximants to be generally useful, we need to show that their degree of contact with the defining series goes to infinity with $\mathcal{M}$.

Theorem 5. *Given an essentially unique, minimal algebraic approximant whose polynomials satisfy $\vec{Q}_{\vec{m}}(0) \neq \vec{0}$ and for which the error in the polynomial defining equations is $O(z^{\mathcal{M}})$, the degree of contact c of the algebraic approximant with the defining series is bounded by*

$$c \geq \left[\frac{\mathcal{M}-1}{\mu}\right], \tag{7}$$

where $[x]$ denotes the greatest interger less than or equal to x, and $1 \leq \mu \leq k+1$ is the smallest value of ν for which the coefficient of $[f(z) - f(0)]^\nu$ in (8) does not vanish for $z = 0$. The value of (7) for $\mu = k+1$ is always a bound.

Proof: Let us expand the polynomial defining equation about the given value of $f(0)$. We get,

$$\sum_{\nu=0}^{k+1} \frac{[f(z)-f(0)]^\nu}{\nu!} \left(\sum_{j=\nu-1}^{k} \frac{Q_{j,m_j}(z)(j+1)![f(0)]^{j+1-\nu}}{(j+1-\nu)!} \right) = O(z^{s+t+1}). \tag{8}$$

Now we note that the $\nu = 0$ term vanishes at $z = 0$ as is required by the polynomial defining equations. If the second term does not vanish at $z = 0$, then the deficiency is zero. It is the case, when we denote the coefficient of $[f(z)-f(0)]^\nu$ by $\mathcal{Q}_\nu(z)$, that for at least one $\nu = \mu$ that $\mathcal{Q}_\mu(0) \neq 0$. This result can be seen as follows. First consider the case $\nu = k+1$. If $\mathcal{Q}_{k+1}(0) \neq 0$, we have the desired result. Otherwise, consider the case, $\nu = k$. In this circumstance, $\mathcal{Q}_k(0) = 0$ if and only if $Q_{k-1,m_{k-1}}(0) = 0$. We can continue in this manner through all $(k+1 \geq \nu > 0)$ the $\mathcal{Q}_\nu(0)$'s. We must eventually find one which does not vanish, since by hypothesis, $\vec{Q}_{\vec{m}}(0) \neq \vec{0}$. In order to focus on the computation of c, it is simplest to consider the determination of f_1. Let μ be the smallest value of ν for which $\mathcal{Q}_\nu(0) \neq 0$. From (8) there will be a non-vanishing term, proportional to f_1^μ, and potentially other terms of order z^μ involving f_1 but no other series coefficients except f_0. Let us now examine the series in z term by term. The first term (beyond the $\nu = 0$ term) is linear in f_1 and if the coefficient does not vanish, we can solve for f_1 directly. If it does vanish, then the second term will be a quadratic equation in f_1 and there will be no mixed terms involving f_1 or terms involving f_j with $j > 1$ because their coefficients vanish as a consequence of the vanishing of the first term. The examination continues in this manner so that we either find an equation for f_1, or the form of the next order term in z involves f_1 alone. We are however guaranteed that we will find an equation for f_1 with at least one non-zero coefficient by the z^μ term. It will be of the μth order in f_1 and corresponds to μ roots which are coincident at the origin. If the appropriate root of this equation for f_1 is not a multiple root, then we can solve for all the rest of the coefficients, f_j, $j > 1$ because they will first appear linearly and if $R(f) = 0$ is the equation for f_1 then their coefficient will always be $\left.\frac{\partial R(f)}{\partial f}\right|_{f_1}$ which is non-zero for this case. By analysis of this case through some rather tedious but straightforward algebra, one can see that there will always be an equation with at least one non-vanishing coefficients to determine f_j which includes the term $\mathcal{Q}_\mu(0) f_j^\mu z^{j\mu}$. Thus the accuracy-through-order equations are always sufficient to determine at least $[(\mathcal{M}-1)/\mu]$ coefficients. ∎

An illustration of these theorems is given by the function

$$f(z) = -1 + z - \frac{1}{2}z^2 + \frac{3}{8}z^3 - \frac{5}{16}z^4 + \frac{35}{128}z^5 - \frac{63}{256}z^6 + O(z^7), \tag{9}$$

which satisfies,

$$(1+z)[f(z)]^2 + 2(1+z)f(z) + 1 + z - z^2 = O(z^7). \tag{10}$$

Condition (5) becomes in this case,

$$2(1+z)(y(z)+1)\frac{dy(z)}{dz} = -[y(z)]^2 - 2y(z) - 1 + 2z + O(z^6). \tag{11}$$

We see by direct substitution of the defining series in (11) that, as z divides (11), the deficiency is $d = 1$ thereby reducing the degree of contact by unity. This deficiency is associated with the the coincidence of two solutions at $z = 0$, and the corresponding vanishing of the first two terms in (8). Thus the solution for y(z) is,

$$y(z) = -1 \pm \sqrt{\frac{z^2}{1+z}} + O(z^6). \tag{12}$$

We now consider the equivalence properties. In this regard it is useful to define the Beckermann (1990) minimal polynomials. First, we need an ordering relation. If $\vec{m} = (m_{-1}, \ldots, m_k)$ is a vector in the index space labeling the algebraic polynomials, then the partial ordering relation $\vec{a} \leq \vec{c}$ means that every component of $\vec{a}$ is less than or equal to the corresponding component of $\vec{c}$. The relation $\vec{a} = \vec{c}$ means that every component of the two vectors is equal, and $\vec{a} < \vec{c}$ means that $\vec{a} \leq \vec{c}$ holds but that $\vec{a} = \vec{c}$ fails. Notice that if $a_i > c_i$ and $a_j < c_j$ for $i \neq j$, then $\vec{a}$ and $\vec{c}$ are incomparable by these partial ordering relations.

Definition 6. *A nontrivial solution of (1) for the algebraic polynomials is called a Beckermann minimal solution if, among all the solutions of (1), there is no other nontrivial solution whose degree is less according to the above given partial ordering relation.*

Notice that a minimal solution according to Definition 2, is also a Beckermann minimal solution, but that there may be Beckermann minimal solutions which are not minimal according to Definition 2. We now give the following results.

Theorem 7. *The statement (i) $f(z)$ is a functional element at $z = 0$ which satisfies*

$$\sum_{j=-1}^{k} Q_{j,m_j}(z)[f(z)]^{j+1} = 0, \tag{13}$$

where $\vec{Q}_{\vec{m}}$ is of true nominal degree and essentially unique, is equivalent to (ii) there exists a functional element $g(z)$ at $z = 0$ for which,

$$\vec{\mathcal{Q}}_{\vec{M}} = \vec{Q}_{\vec{m}}, \quad \forall \vec{M} \geq \vec{m}, \tag{14}$$

where $\vec{Q}_{\vec{m}}$ is minimal and $\vec{\mathcal{Q}}_{\vec{M}}$ is a Beckermann minimal solution of type $\vec{M}$ for $g(z)$, the inequality is in the sense of Definition 6, and $g(z) = f(z)$.

Proof: First, (i) implies (ii) as, if we pick $g(z) = f(z)$, then $\vec{Q}$ is a solution of type $\vec{M}$ for any $\vec{M} \geq \vec{m}$ and as by hypothesis, as $\vec{Q}_{\vec{m}}$ is minimal, it is at least a Beckermann minimal solution of type $\vec{M}$.

Second we must consider, whether (ii) implies (i). Suppose that (ii) holds, therefore we must have

$$Q_{-1,m_{-1}}(z) + \sum_{j=0}^{k} Q_{j,m_j}(z)[g(z)]^{j+1} = O(z^t), \quad \forall t < \infty. \tag{15}$$

By the minimality of $\vec{Q}_{\vec{m}}(z)$, there are no common factors of the type $1 + az$ and the division by any factor of z^j leaves (15) unchanged except that $m_i \to m_i - j$, as it holds for all $t < \infty$. As we have seen in the proof of Theorem 4 that we can divide both sides of (5) by z^d, the integration of this divided equation shows by the implicit function theorem that $g(z)$ is a functional element. Thus the left-hand side of (15) is regular at $z = 0$ and so by (15) must be identically zero. The remaining problem to complete the proof is to show that $g(z) = f(z)$. This result follows from the accuracy-through-order Theorem 4 or 5 and the principle of analytic continuation. The point is that either of these theorems show that the degree of contact at the origin between the algebraic approximant and $g(z)$ is infinite, provided that at most $d + 1$, where d is the deficiency, of the initial conditions $y^{(j)}(0) = g^{(j)}(0)$ hold. This result, by the principle of analytic continuation, proves that any $y(z) = g(z)$ and so $f(z) = g(z)$. ∎

Acknowledgements

I wish to thank P. R. Graves-Morris for helpful discussions. I also wish to thank the referee for pointing out an incompleteness in the original version.

References

Baker, George A. Jr.: 1984, 'Invariance properties in Hermite–Padé approximation theory', *J. Comp. & Appl. Math.* Vol. no. **11**, pp. 49–55

Baker, George A. Jr. and Graves-Morris, P. R.: 1990, 'Definition and uniqueness of integral approximants', *J. Comp. & Appl. Math.* Vol. no. **31**, pp. 357–372

Baker, George A. Jr. and Graves-Morris, P. R.: 1994, 'Existence of certain sequences of Hermite-Padé approximants', *J. Comp. & Appl. Math.* to be published

Beckermann, B.: 1990, 'The structure of the singular solution table of the M-Padé approximation problem', *J. Comp. & Appl. Math.* Vol. no. **32**, pp. 3–15

McInnes, A. W.: 1992, 'Existence and uniqueness of algebraic function approximations', *Constructive Approximation* Vol. no. **8**, pp. 1–21

Padé, H.: 1892, 'Sur la répresentation approchée d'une fonction par des fractions rationelles', *Ann. de l'Ecole Normale Sup.* $3^{iéme}$ *Série* Vol. no. **9**, Suppl., pp. 3–93

Shafer, R. E.: 1974, 'On quadratic approximation', *SIAM J. Num. Anal.* Vol. no. **11**, pp. 447-460